Project portfolio management at Northwestern University Libraries: Developing an innovative solution to managing project life cycles

Received (in revised form): 25th August, 2017

Carolyn Caizzi
has worked at Northwestern University Libraries for five years and is Head of Repository and Digital Curation, a department of 17 staff specialising in digitisation, open source software development, digital object curation and metadata creation. She previously worked at Yale University Library and the University of Denver. She earned her BA from Brown University, MFA from University of Illinois at Chicago and MSLIS at University of Illinois at Urbana-Champaign.

Northwestern University Libraries, 1970 Campus Dr., Evanston, IL 60208, USA
Tel: +1 847 467 3898; E-mail: carolyn.caizzi@northwestern.edu

Jennifer B. Young
is the Metadata Coordinator at Northwestern University Libraries, where she manages the creation of metadata for the digital repositories. She previously worked at Saint Louis University and has been at Northwestern for 14 years. She earned her BA from Bradley University and her MSLIS from the University of Illinois at Urbana-Champaign.

Northwestern University Libraries, 1970 Campus Dr., Evanston, IL 60208, USA
Tel: +1 847 491 8978; E-mail: j-young2@northwestern.edu

Dan Zellner
is the Production Coordinator at Northwestern University Libraries. He has been involved in managing digitisation workflows and projects for over 15 years. He is also currently working with the Digital Library Federation's Cost Assessment group in the development of the Digitisation Cost Calculator.

Northwestern University Libraries, 1970 Campus Dr., Evanston, IL 60208, USA
Tel: +1 847 467 1399; E-mail: d-zellner@northwestern.edu

Abstract The Northwestern University Libraries Repository and Digital Curation (RDC) department is tasked with digitising and creating digital collections from the materials managed by the Libraries' curators. The previous process for the intake and management of these projects created a situation of feast or famine regarding project work that was ready and vetted for staff to begin. With the development of a new process, the RDC points process, the department is now better positioned to manage its portfolio of projects. This paper provides details concerning the problem and the methods used to develop a new process to address it. The paper outlines the new process of project and proposal intake, determination of capacity (available staff work hours), time-tracking and scheduling. The paper also cites the sources that inspired the RDC points process and aims to serve the same purpose for others faced with similar project portfolio management challenges.

KEYWORDS: project management, digitisation, workflows, collaboration

INTRODUCTION

Northwestern University Libraries (NUL) supports the university's mission of discovering and creating new knowledge by providing access to research collections of international distinction and delivering advanced scholarly digital information services anytime, anyplace. NUL contains over 6.7 million volumes, 24,757 linear feet of archives and manuscripts, access to over 201,000 electronic journals, and has digitised or acquired over 275 TB of unique scholarly and archival digital content. As an organisation, NUL has over 200 staff involved in acquiring, describing, making accessible and preserving the collections as well as supporting faculty and students with their research, learning and teaching needs. In 2014–15, the Libraries underwent a massive restructuring to improve processes in order to support faculty, students and outside researchers more efficiently. One of the newly formed departments is the Repository and Digital Curation (RDC) department. RDC is responsible for digitisation, metadata creation and the development of repository software to preserve and make accessible library collections and scholarly works. As part of workflow redesign, RDC needed to improve the process to deliver library digital collections in a consistent and timely manner for faculty and outside researchers. These digital collections consist of items such as archival ethnographic photographs, maps, posters, postcards, audio recordings, film, video and texts. Collections are proposed by a library curator as a digital collection project for RDC to manage. At present, there are six curators at the Libraries who represent the needs of researchers and faculty on campus.

The creation of these digital collections is a project in and of itself. For example, the work that goes into creating a collection involves reviewing the intellectual property rights of the collection, a conservation review, as well as inventorying the physical items to be digitised. The digitisation team members produce high-quality captures of the photos, texts, film, video or audio that comprise the collection. Metadata specialists enter descriptions of each image or audiovisual object, so that end users — the researchers and faculty — can easily find it amid hundreds of thousands of digital objects. Quality control of digital objects and associated metadata is a necessary part of the process, allowing for the ingest of the objects into the software that helps preserve these assets and make them accessible either to the world or to the university community. These are just some of the major activities that go into making a digital collection from special and archival collections within an academic research library setting.

Prior to the restructuring of the Libraries, the essential problem with the digital collection creation process was that the selection of projects to be worked on was divorced from both the curators and the staff who would be performing the work. Additionally, the staff most affected by the selection of projects needed a better way to manage a portfolio of them. Before restructuring, a digital projects committee voted on the merit of digital collection proposals rather than whether there were sufficient resources to complete the project. The committee also was made up of people from across the Libraries rather than the staff closest to the work. Additionally, the committee met only when there was a proposal to review, and did not have a scheduled call for proposals. This process also put the onus on the curators to write a proposal for a digital collection project. The staff coordinating the digital collection work were either waiting for the next approved project or trying to balance too many projects because of the committee selection process, thereby perpetuating a feast or famine approach to managing a portfolio of projects. In addition, curators would bypass the committee altogether and go straight to the staff responsible for most of the digital collection work, forcing them to prioritise work on a first come,

first served basis. This process resulted in building digital collections that were not always in line with the strategic goals of the Libraries. The restructuring, which dissolved the digital projects committee, allowed RDC to operationalise the work of selecting and managing these projects using a new process, which will now be described.

THE DEVELOPMENT OF A NEW PROCESS

The identified problem, and its associated issues, is certainly not unique to Northwestern. Given this situation, the next step was to investigate what others were doing both internally and externally to see if these approaches could be adopted or developed to address the problem.

Two primary sources of inspiration provided an ideal framework for the solution that was developed. The internal source was the Northwestern University Libraries Preservation Department. The department uses a 'point system' based on the time management system developed at the Library of Congress.[1] The other source was Emory University Libraries' project selection process. These two sources will be detailed in the paragraphs that follow.

Northwestern Libraries' Preservation Department points process has been in place for many years and is used to determine the capacity of its staff and compare this against the work requested. A point is defined as one hour of work. The following formula is used as a starting point in determining the capacity of each staff member:

37.5 hours × 52 weeks – (15 vacation days + 15 sick days +10 holidays per year) = 1,650 hours per year
Total hours at 80 per cent efficiency = 1,320 hours

Once the final determination of available points is established, they are divided between the curators who are the primary users of preservation services in the library. There are two times during the year in which curators submit their requests and these are evaluated in connection with available points for the given six-month period. These request periods involve one-on-one meetings with the curators so that work can be discussed in detail and examples of collection material can be submitted. It is important to mention that the scope of these meetings is limited to preservation work only.

In contrast, Emory University Libraries takes a comprehensive approach towards project intake. As Emory Libraries Digital Collection and Development Policy states: 'in order to document digitisation efforts and plan work and manage capacity, all digitisation and preservation efforts should come through and be reviewed by the Digital Collections Steering Committee and relevant Subcommittee'.[2] This approach involves proposal submission that is evaluated and reviewed by a committee that represents the various parties involved in the work or its oversight.

The Preservation Department's approach to managing its work was very attractive. The one-on-one interviews with customers (ie the curators) makes the process of solicitation and intake of work more approachable and more likely to move forward than a formal proposal submission process. The negative aspect concerning this approach is that the focus of the process is very limited: it addresses preservation work solely. This narrow focus presents the risk of missing dependencies and interrelated elements that may affect the work. Given the especially interrelated and collaborative nature of digital collection work, the Preservation Department's approach could not simply be adopted.

Emory University Libraries' approach is a more holistic way to conduct the intake and vetting of projects. The process involves a steering committee and a subcommittee comprised of stakeholders and also key members of the teams that do the work.

Proposal submission is a required part of this system; this is probably due to the number of people involved in the evaluation process. Northwestern's previous attempt at committee structures and proposal requirements had not fared well so this system could not be adopted outright. Many of the elements of this system, however, represented ways that committees and groups can be utilised effectively in creating and managing a portfolio of projects.

Both the Northwestern University Libraries Preservation Department's and the Emory University Libraries' processes are well-suited for their respective department and institution. The challenge was to synthesise these two approaches to create a process for the RDC Department that is nimble yet addresses the interconnection between groups doing related work on a project and those affected by the work. The RDC process that was developed, and continues to be refined, attempts to capture these qualities.

RDC POINTS PROCESS OVERVIEW

The process consists of four primary components:

- proactive proposal gathering;
- work capacity determination;
- customer-based project selection based on capacity;
- scheduling based on customer needs and resource availability.

Given the four components, RDC had to delineate an annual project cycle. The annual calendar for the process is as follows:

- *Spring quarter (March–May):*
  - quarterly meeting: final presentation of projects to curators
  - determine capacity for next project cycle
  - finalise list of projects and schedule for new project cycle
- *Summer quarter (June–August):*
  - quarterly meeting — formally present the project schedule
  - project work
- *Autumn quarter (September–November):*
  - quarterly meeting — progress update
  - project work
- *Winter quarter (December–February):*
  - quarterly meeting — progress update
  - wrap up project work and close out projects
  - curator interviews for next project cycle in February

This annual process was designed to align with other organisational activities and practices. Northwestern University follows a quarter system, so the project process cycle is likewise divided into quarters. Additionally, the university-wide annual staff evaluation process runs from March to February of the following year, so it was essential to have the digital project selection and work process start and end at the same time as the evaluation process. The fiscal year for the university, however, starts in September and goes to August of the following year. In this case, it was beneficial to have the project cycle begin before the fiscal year in order to assess funding needs for the new fiscal year (ie new equipment, additional personnel, etc). With the schedule presented, the four components of the process can be detailed.

PROACTIVE PROPOSAL GATHERING

The process begins with proactive proposal gathering. Rather than having customers (the curators) submit a formal proposal, one-on-one meetings are scheduled with each curator at the beginning of the project cycle. The curators are responsible for simply organising their requests, providing some of the details (sample materials for digitisation, etc) and, most importantly, prioritising their requests. A structured interview takes place in which basic information about the projects is captured. This new process was

much welcomed by the curators and did not put the onus on curators to write a formal proposal. The process also allows for dialogue, which saves a good deal of time when clarification is needed regarding the details of a project. Following the meetings, a first-level estimate is generated based on the number of items in the project and existing metrics. This estimate gives a general idea of the scale and scope of the project: a more in-depth estimate is generated for projects that have made the shortlist for the project cycle.

WORK CAPACITY

One of the most challenging parts of the process is determining the capacity of the RDC department for a given year. The formula mentioned earlier is used to determine a staff member's total available working hours. The efficiency rating of 80 per cent (ie 80 per cent of a staff member's time is available for work) is cited frequently as a guideline used by industrial engineers. While this seemed a reasonable starting point for RDC capacity measurement, project management expert Joseph Heagney notes that:

> the only people who are available to do work 80 per cent of the time are those whose jobs tie them to their workstations. This is true for factory workers and others who do routine work (and even these people move around). With knowledge workers you never get 80 per cent of a day in productive work. The figure is usually close to 50 per cent, and it may be lower.[3]

Additional time is taken off each staff member's capacity for other non-project–related work, resulting in variation between the final work capacities for each staff member. At this point in the development of the process, a more detailed utilisation analysis has not been done. This would involve an investigation into periods of the year and general availability of staff. For now, a general expression of capacity is used for the availability of staff to do work during the annual project cycle.

A staff member's time is broken down by activities (project and non-project related) and expressed in percentages: percentage of time spent in meetings, supervision, etc. Three primary categories of activity are also quantified and especially important for digital project work: digitisation, metadata and project management. Capacity for digital project work is determined by these three categories for each staff member. Once the entire department's capacity for digitisation, metadata and project management work is determined, 60 per cent of this capacity is allocated to projects that come out of the RDC points process and 40 per cent of capacity is reserved for *ad hoc* projects that come up during the year. *Ad hoc* projects are typically distinctive collection requests in support of faculty research and other needs that may not be anticipated in advance.

CUSTOMER-BASED PROJECT SELECTION BASED ON CAPACITY

Once all the interviews have been concluded and the high-priority projects are identified, a list of these high-priority projects, along with corresponding work estimates, is presented to the key administrators who have strategic oversight of the Libraries' special and archival collections. The estimate for each project states the total digitisation, metadata and project management hours required for its completion. Administrators are also given the total capacity of the RDC department in terms of digitisation, metadata and project management points along with any available funding to hire a vendor or additional staff to do work. Using the RDC capacity, the funding and the estimate for each high-priority project, the stakeholders come together to draft their shortlist of projects for the next project cycle. This way of working allows the internal customers to choose projects strategically across curatorial

divides and increases transparency. Once the shortlist is established, a deep-level assessment of all the shortlist projects takes place to make sure that no factors involving the work for these projects have been missed. If the deep assessment identifies problem areas for certain projects then the estimate needs to be adjusted and the list revised. On occasion, projects are phased in order to accommodate curator needs or to address capacity limitations.

SCHEDULING BASED ON CUSTOMER NEEDS AND RESOURCE AVAILABILITY

After the project list has been finalised, the schedule is created. A schedule that addresses customer (curator) needs and their deadlines is a priority. In conjunction with those considerations is resource availability (staff, equipment and funding). As much as possible, related work is grouped together and timed in order to make the best use of staff time. As mentioned earlier, an analysis of resource utilisation based on time of year has not been done but there are certainly known factors, such as the holiday season, vacations and major system upgrades that factor into planning for the project cycle.

IMPLEMENTATION OF THE RDC POINTS PROCESS

The implementation of any new process naturally involves a commitment of time and resources. One of the first steps was an assessment of the department or unit doing the work. In the case of RDC, this involved a holistic look at the department's processes as well as the processes of related departments, then identifying dependencies and points where work and actions in one department affect another. Beyond the department, existing administrative structures (the academic calendar, the fiscal year and the employee evaluation schedule) also had to be factored into the design of the RDC points process and its corresponding schedule. The utilisation of existing structures provided a point of reference for the new RDC points process and assurance that the process was not an isolated entity with no relation to any of the work and activities that take place in the library.

To implement the RDC points process, an investigation of RDC work and time-tracking systems was also required. Previously, project work was tracked using Atlassian's Jira software, but the cost and also the utility of the application presented difficulties that prompted the need to move to a new system. After some investigation, Smartsheet was adopted due to its cost and ease of use; another point in its favour was its campus-wide adoption by Northwestern University. Previously, RDC was not in the practice of time-tracking. Work was timed in order to generate metrics but tracking time spent on project work was simply not done. Initially, a web-based program called Toggl was tested, but while it was well designed and proved easy to use, the cost was prohibitive for the number of users required to track time. Smartsheet was again used to track project time and, on the positive side, provided an application to solve work and time-tracking. Consistent time-tracking continues to be a challenge within the department but is fundamental in order to properly assess the time required for any given task or project.

Reporting is inherent in the activity of work and time-tracking. Reports are generated throughout the process for internal and external audiences and are essential. Unfortunately, Smartsheet reporting is not as robust as a database application such as Microsoft Access but views of work and lists can be generated to keep staff, project managers and curators informed. Likewise, process documentation is critical and something as simple as agreeing on where to store documentation was absolutely necessary in order to ensure changes to the process are noted and remembered.

The rollout of the RDC points process required multiple announcements and meetings to educate all regarding its implications and expectations. Time-tracking, for example, needed to be communicated to staff, and while it seemed to them invasive at first, it was later seen to be helpful in illustrating the 'cost' of a project. Customer, team and stakeholder assessment is also a critical part of the process. Quarterly project meetings with curators function to keep curators informed regarding project progress and also to remind them of the process and how it works. Project teams as well as project managers meet to check on progress and address any problems or issues concerning the work.

In addition to meetings with both stakeholders and project teams, time must be set aside to assess how the process itself is working. The following section assesses the process since its implementation two years ago, citing both its benefits and areas for further development.

BENEFITS

An immediate benefit of moving to this process is that RDC knew what projects were on deck for the upcoming year. It allowed RDC to better plan its work for its staff and finally move towards breaking the feast or famine cycle. Previously, the Libraries' projects came from multiple places and at any time of year. The Libraries' projects committee process could take up to a year to move from proposal to the start of work. Additionally, the curators and other library stakeholders would often bypass the committee entirely and drop off projects, usually with a quick turnaround time. All of this made it difficult to report about what projects were being worked on at any given point in time during a fiscal year. During times of feast, RDC was overworked and overwhelmed. During times of famine, RDC had to devise work to do. This unpredictability also made it difficult to report out for annual report and employee reviews. It was impossible to say which projects were finished or where they were in the project cycle as the cycle did not exist. RDC had disparate tracking mechanisms in place with no one-stop shop to quickly determine status of a project or even update those statuses. Now, RDC plans the entire project from the beginning, including the final deliverable. Final deliverable options include submission to HathiTrust, ingestion into the local repository and digital collections websites or as part of an online exhibit site. The project proposer decides which deliverables are needed at the time of the proposal instead of later in the process. Prior to the implementation of the RDC points process, there was sometimes a lack of decision or agreement about the final deliverable for a project at its start, which led to confusion and project stalling.

As seen in Figure 1, in 2016 RDC had 35 projects in its project cycle. RDC carried over 11 legacy projects that had not yet been completed at the start of this new process. RDC received 15 new scheduled projects for the year and took on 9 *ad hoc* projects as the year progressed. Discounting the legacy projects, RDC managed to come close to its original planned division of capacity (ie 40 per cent capacity dedicated to *ad hoc* projects and 60 per cent to scheduled projects). In the 2016 Project Cycle, 38 per cent of capacity was spent on *ad hoc* projects and 62 per cent spent on scheduled projects. By the end of the project cycle, RDC completed 9 of 11 legacy projects, 6 of 9 *ad hoc* projects and 9 of 15 scheduled projects.

COMMUNITY OF TRUST

Before the reorganisation of the Libraries, RDC's work was much less transparent to the rest of the Libraries. The Libraries digital projects committee consisted of people who were mostly removed from the actual work needed to complete these projects. One reason for the reorganisation was to foster culture change, including pushing

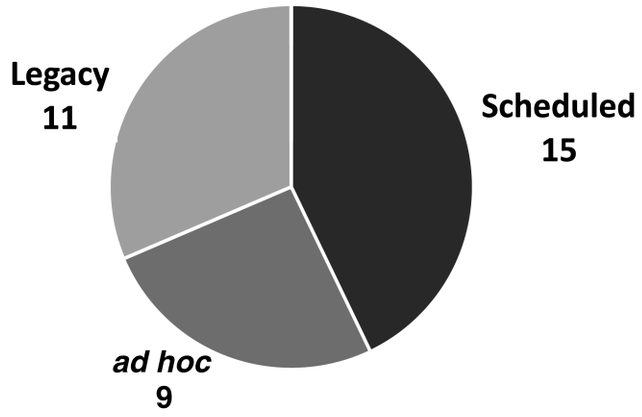

Figure 1: 2016 project cycle

decision-making power to the appropriate level of staff. Not having the committee allows the curators to see the proposed projects and schedule. This transparency also allows them to see what their peers are proposing and possibly influence their own projects. As part of RDC's project cycle, quarterly meetings are held with the curators for progress reports as well as to talk about the process in general. There is much more collaboration built into the process from the beginning. Lastly, having a yearly cycle means status reports are now easily generated and the information about projects remains fresh.

EFFICIENCIES GAINED

Before RDC had this process, it was impossible to anticipate when the digital projects committee would select new projects. As there was no project selection cycle, the capacity of RDC was always in flux. Additionally, it was impossible to anticipate when a curator would bypass the committee. The old process was very opaque to the staff managing the work, the staff doing the work as well as the stakeholders. Instituting the points process has given RDC a better understanding of what it can handle in a given year and enabled it to manage the portfolio of projects much more efficiently. RDC can anticipate productivity for a project cycle and can create benchmarks and goals for future cycles based on data.

Figure 2 shows three different projects RDC worked on during its last project cycle. For the scheduled project, 2016 Football Films, and an *ad hoc* project, Turkish Oral History, the actual staff time spent was pretty close to the initial estimates. For the scheduled project, Africa Embraces Obama 3D, however, the estimates were very off compared with the actual time spent. One reason was the unanticipated conservation work needed as part of this project. Such work needs to happen before the digitisation of an object. The initial inventory of the objects provided by the curatorial staff provided starting information for both the shot sheet for digitisation and the creation of metadata. Large gaps of information about these three-dimensional objects made it more time-consuming to compare the initial inventory with the actual pieces. While seeing a project's estimates being off is not ideal, it was immensely useful information for RDC. Without the initial estimates and tracking systems, RDC would have had no idea how much actual time it was spending on this project. In fact, RDC decided to phase this project over multiple years after getting a better idea of the time needed to complete it. This was done after consulting with the curator to get her input as well.

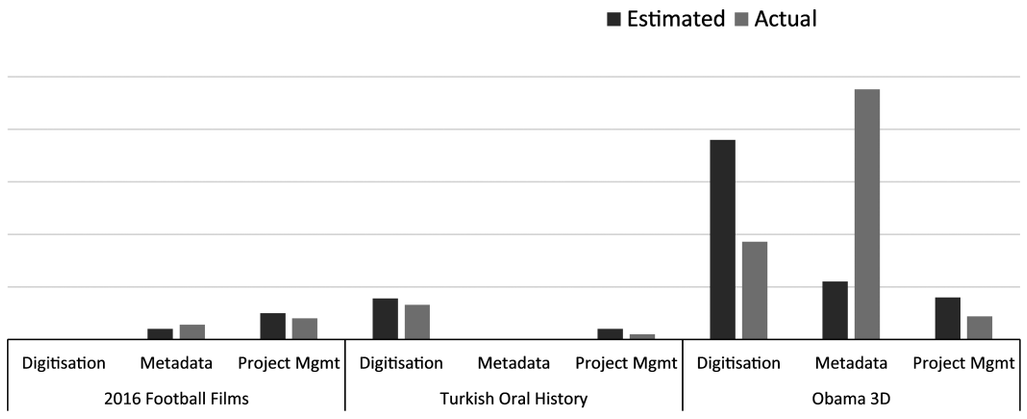

Figure 2: Estimate versus actual time

RDC continually assesses projects throughout the project cycle. As mentioned earlier, RDC meets regularly with the curators and other stakeholders. Additionally, the project managers meet weekly to discuss the projects and the process. Communication outside of these times happens as often as needed. Having these channels open for discussion also allows the process to be more iterative, for example, recognising that a project would be better served by being phased over several years instead of attempting to do it in a single burst. These discussions also help eliminate scope creep.

CHALLENGES

A particular challenge faced by RDC during the implementation of this process was the loss of two project managers, meaning the remaining ones had to pick up more projects until replacements could be hired. Thirty-five projects had to be divided among three project managers from a total of five. Staff turnover will always be a challenge for planning based on capacity, but over time, the turnover rate could be factored into capacity to help address the challenge.

Another challenge was the transition to recording staff time spent on different aspects of the projects. Time-tracking was new for many staff and RDC needs to be better about tracking as a whole. To gain participation, RDC communicated to staff that the goal of time-tracking is not to be punitive, but to have better data for initial project estimates so that projects are completed. The project managers have demonstrated that they use this information both for planning purposes as well as post-project assessment and all staff now track their time more accurately.

A final challenge was keeping documentation current and consistent. During this timeframe, Northwestern University and the Libraries licensed new tools to be used among staff — in particular, Sharepoint, Box and Smartsheet. Previously, RDC's project documentation was spread out among an intranet, Jira, Google Drive and documents on shared internal drives. With the move to new tools, RDC is slowly making documentation consistent and current. RDC also did not have a formalised set of steps to follow for each project as each project manager had worked in silos prior to the restructuring. As part of reorganising and combining its project documentation, RDC also created a checklist template for project managers to follow for the different types of digital collection projects.

CONCLUSION

In developing a process based on the solutions created by others, RDC has remedied the situation of feast and famine regarding project work that was

ready and vetted for staff to begin. It is important to mention that RDC's past methods recorded in this paper reflected an earlier period in the department's history when there were fewer staff and fewer projects — hence the need for portfolio management was less pressing. What often gets lost when embarking on projects is the need for a process for managing them. It is also equally important to note that the development of a process for managing projects is a project in and of itself. The RDC points process has been in effect for two years. It will continue to be adjusted and adapted as workflows and organisational structures change.

REFERENCES

1. Harris, K.E. and Schur, S.E. (2006) 'A Brief History of Preservation and Conservation at the Library of Congress', Library of Congress, Preservation Directorate, Washington, DC, available at: https://www.loc.gov/preservation/about/history/pres-hist.pdf (accessed 2nd June, 2017).
2. Robert W. Woodruff Library, Emory. (n.d.) 'Digital collection development and preservation policy', available at: http://guides.main.library.emory.edu/ld.php?content_id=6155258 (accessed 2nd June, 2017).
3. Heagney J. (2012) 'Fundamentals of Project Management', 4th edn, American Management Association, New York, NY.